**Cowles Foundation Paper 832**

*Econometrica*, Vol. 61, No. 1 (January, 1993), 139–165

# EXACTLY MEDIAN-UNBIASED ESTIMATION OF FIRST ORDER AUTOREGRESSIVE / UNIT ROOT MODELS

BY DONALD W. K. ANDREWS[1]

This paper is concerned with the estimation of first-order autoregressive/unit root models with independent identically distributed normal errors. The models considered include those without an intercept, those with an intercept, and those with an intercept and time trend. The autoregressive (AR) parameter $\alpha$ is allowed to lie in the interval $(-1, 1]$, which includes the case of a unit root. Exactly median-unbiased estimators of the AR parameter $\alpha$ are proposed. Exact confidence intervals for this parameter are introduced. Corresponding exactly median-unbiased estimators and exact confidence intervals are also provided for the impulse response function, the cumulative impulse response, and the half life of a unit shock. An unbiased model selection procedure is discussed. The procedures that are introduced are applied to several data series including real exchange rates, the velocity of money, and industrial production.

KEYWORDS: Autoregressive process, confidence interval, deterministic time trend, impulse response function, median-unbiased estimation, model selection procedure, stochastic time trend, unbiasedness, unit root model.

## 1. INTRODUCTION

THIS PAPER CONSIDERS POINT and interval estimation in a first-order autoregressive/unit root (AR/UR) model that may contain an intercept and time trend and has independent identically distributed normal errors. Exactly median-unbiased estimators and exact confidence intervals are introduced for the AR parameter, the impulse response (IR) function, the cumulative impulse response (CIR), and the half life of a unit shock (HLS). An unbiased model selection procedure is also introduced.

A basic motivation for this paper is the emphasis placed in the unit root literature on testing for a unit root and the relative neglect of point and interval estimation. A problem that arises with hypothesis tests in the unit root context is that tests have low power in many scenarios of empirical interest (e.g., see DeJong, Nankervis, Savin, and Whiteman (1992a, b)). In such cases, the failure to reject the null hypothesis cannot be construed as providing evidence in favor of the null.[2] One needs to provide additional information. Point and interval estimators are statistics that can be used for doing so.

[1] I would like to thank Hong-Yuan Chen, Eric Zivot, and Zhaoliang Zhu for their excellent research assistance. I would also like to thank two referees, a co-editor, David DeJong, Jean-Marie Dufour, Eric Ghysels, Pierre Perron, Peter Phillips, Glenn Rudebusch, Chris Sims, Bent Sorensen, Jim Stock, and participants of the finance workshop at the Olin School of Business, Washington University for helpful comments, and Willem Buiter for supplying the exchange rate data analyzed in Section 9. I gratefully acknowledge the research support of the National Science Foundation via Grant No. SES-8821021. I dedicate this paper to my wife Anne Kao Andrews whose continual support is responsible for any and all research success I have had to date (and who has promised unspeakable repercussions if I do not dedicate the paper thus).

[2] See Andrews (1989) for further discussion of this problem in the context of nonlinear models without trends.

The problem with utilizing estimators in AR/UR models is that of bias. Standard estimators, such as the least squares (LS) estimator, are significantly downward biased in AR/UR models that contain an intercept or an intercept and time trend, especially when the AR parameter $\alpha$ is large. For example, in the model with intercept and time trend, sample size equal to 60, and normal errors, the LS estimator has downward median-biases when the AR parameter $\alpha$ equals .7, .85, and 1.0 of .08, .09, and .15 respectively (where median-bias is the difference between the median of the estimator and the true parameter value). Even when $\alpha$ is as small as .3, the downward bias is .05. These downward biases are considerable, especially when $\alpha$ is large. They cause the LS estimator to be a misleading indicator of the true value of $\alpha$. For example, the probabilities that the LS estimator underestimates the AR parameter $\alpha$ in the model above when $\alpha$ equals .3, .7, .85, and 1.0, are .66, .78, .87, and .995 respectively. (Note that in models with no intercept or time trend, which are of relatively limited practical importance, the downward bias of the LS estimator is very much smaller than it is in models with an intercept or with an intercept and time trend. For models with no intercept, the LS estimator is not a misleading indicator of the true value of $\alpha$, at least for samples of size sixty or more.)

To deal with the problem of downward bias, this paper introduces an exact bias correction for the LS estimator. The method is as follows: If the LS estimate equals .8, say, one does not use .8 as the estimate of $\alpha$, but rather, one uses the value of $\alpha$ that yields the LS estimator to have a median of .8. In the model described above this is .9. With this bias correction, the estimator is exactly median-unbiased. Note that the magnitude of the correction can be quite large—a LS estimate of .85 corresponds to a median-unbiased estimate of 1.0 in the model described above.

A similar procedure can be used to obtain exact confidence intervals for $\alpha$. These confidence intervals are useful in their own right, as indicators of the variability of the median-unbiased estimator, or as the basis of exact hypothesis tests. We focus attention on *central* confidence intervals which by definition have equal probabilities of overestimation and underestimation.

*Median-unbiased* estimators and *central* confidence intervals are emphasized in this paper because of their impartiality properties. In cases where the magnitude of a parameter is a contentious issue, as it is in (trend) stationary versus unit root debates, it is useful to have statistical procedures available that treat different parameter values on an even footing. Classical hypothesis tests do not do this and neither do proper Bayes estimators.

It is interesting to note that much of the work on Bayesian estimation of AR/UR models concerns estimation using some version of a flat or noninformative prior. For example, see DeJong and Whiteman (1991), Schotman and van Dijk (1991a, b), Sims (1988), and Sims and Uhlig (1991) for the use of a flat prior and Phillips (1991) for the use of a noninformative prior. The motivation for the use of such priors is the desire for a certain degree of impartiality. This is similar to the motivation for using median-unbiased estimators and central confidence intervals. The latter two, however, avoid the difficult question of

choosing the most appropriate prior to use. They also avoid the problems that arise due to the question of whether the prior used is really noninformative; e.g., see the discussion of Phillips (1991). On the other hand, Bayes estimators do possess other attributes that are attractive; see the references above for a discussion of these.

The point and interval estimation procedures considered here have the following attributes. They are exact procedures rather than asymptotic ones. They exhibit a smooth transition between the (trend) stationary case when $|\alpha| < 1$ and the unit root (random walk) case when $\alpha = 1$ (in models with an intercept or an intercept and time trend). They utilize desirable assumptions regarding initial conditions—if $|\alpha| < 1$, a stationary initial condition is imposed and if $\alpha = 1$ an arbitrary initial condition is allowed. In addition, the median-unbiased estimator is unbiased whether the AR parameter $\alpha$ is fixed or is random, as in a Bayesian context. In the latter case, it is median-unbiased for all possible distributions of $\alpha$.

A drawback of the exact procedures considered here is that they only apply to first-order AR processes and not to more general $p$th order processes. The extension of the procedures of this paper to AR($p$) models is considered in Andrews and Chen (1992); see also Rudebusch (1992) and Stock (1991). For AR($p$) models, the procedures are no longer exact; they are only approximate.

A second drawback of the procedures considered here is that they require the specification of the distribution of the innovations in the AR/UR model. If the distribution is specified incorrectly, then the procedures will not be exact. On the other hand, numerical results presented below show that the procedures based on normally distributed innovations are surprisingly robust against a variety of nonnormal distributions. For such distributions, the procedures are approximately valid and the approximation error is small. The median-unbiasedness of the estimators considered is also found to be quite robust against ARCH innovations. The confidence intervals considered, however, are not as robust against ARCH innovations.

A third drawback of the median-unbiased estimator that is considered here is the lack of an explicit optimality property for it—we do not know whether it is a best median-unbiased estimator. It is possible that the estimator does not fully exploit all the information in the sufficient statistics for the parameters. Note that the LS estimator, which is widely used in practice, is also subject to the latter criticism.

There are a number of papers in the literature that are related to the present one. For stationary AR processes, Quenouille (1949, 1956) introduced the jackknife estimator of the AR parameter $\alpha$ that is mean-unbiased to order $1/T$ as $T \to \infty$. His results do not allow for a time trend or for an AR parameter of one. Hurwicz (1950), Marriott and Pope (1954), and Kendall (1954) established the mean-bias of the LS estimator in the same model as considered by Quenouille (1949). Orcutt and Winokur (1969) used the latter results to construct approximately mean-unbiased estimates of the AR parameter in stationary models. Le Breton and Pham (1989) calculated the exact and asymptotic

biases of the LS estimator in stationary, unit root, and explosive AR(1) models without intercepts. Shaman and Stine (1988) and Stine and Shaman (1989) established the mean-bias of the LS estimator to order $1/T$ in stationary AR($p$) models. Dufour (1990) introduced exact confidence intervals (based on the Durbin-Watson statistic) for the autoregressive coefficient in a linear regression model with AR(1) errors. Independent of the present paper, Rudebusch (1992) considered approximately median-unbiased estimators of the AR parameters in an AR($p$) model with a time trend and Stock (1991) introduced asymptotic confidence intervals for the largest root in an AR($p$) model with a time trend (using largest root local to unity asymptotics). Lastly, we note that there are numerous papers that consider hypothesis tests in the AR/UR models considered here. For brevity, we do not provide references.

The remainder of this paper is outlined as follows. Section 2 defines the models that are considered and points out a useful invariance property of the LS estimator for these models. Section 3 introduces the median-unbiased estimator of the AR parameter $\alpha$. Section 4 provides exact confidence intervals for $\alpha$. Section 5 discusses median-unbiased estimators and exact confidence intervals for the IR function, the CIR, and the HLS. Section 6 introduces an unbiased model selection procedure. Section 7 investigates the robustness of the procedures introduced in Sections 2–6 to some forms of nonnormality and conditional heteroskedasticity of the innovations in the AR/UR models. Section 8 provides a discussion of some properties of the LS estimator and Bayes estimators. Section 9 describes the application of the results of the paper to a number of exchange rate series (analyzed previously by Schotman and van Dijk (1991a)) and the velocity and industrial production series of Nelson and Plosser (1982). Appendix A provides a proof of the invariance property stated in Section 2. Appendix B describes the numerical methods that are used to construct the tables given in Section 3.

## 2. DEFINITION OF MODELS

Three models of the time series $\{Y_t\colon t = 0,\ldots,T\}$ are considered. Each is based on a latent AR(1) time series $\{Y_t^*\colon t = 0,\ldots,T\}$:

$$(2.1)\qquad Y_t^* = \alpha Y_{t-1}^* + U_t \quad \text{for} \quad t = 1,\ldots,T, \qquad \text{where} \quad \alpha \in (-1,1],$$

$$U_t \sim \text{iid } N(0,\sigma^2) \quad \text{for} \quad \sigma^2 > 0,$$

$$Y_0^* \sim N\big(0,\sigma^2/(1-\alpha^2)\big) \qquad \text{if} \quad \alpha \in (-1,1),$$

and $Y_0^*$ is an arbitrary constant or random variable if $\alpha = 1$.

As defined, $\{Y_t^*\}$ is a strictly stationary, mean zero, normal, AR(1) process if $\alpha \in (-1,1)$. If $\alpha = 1$, on the other hand, $\{Y_t^*\}$ is a normal random walk with arbitrary initial condition.

The three models for $\{Y_t\}$ are defined as follows:

$$
\begin{aligned}
(2.2)\quad &\text{Model 1:}\quad Y_t = Y_t^* \quad \text{for} \quad t = 0,\ldots,T \quad \text{and} \quad \alpha \in (-1,1),\\
&\text{Model 2:}\quad Y_t = \mu + Y_t^* \quad \text{for} \quad t = 0,\ldots,T,\\
&\qquad\qquad \text{for some} \quad \mu \in R \quad \text{and} \quad \alpha \in (-1,1],\\
&\text{Model 3:}\quad Y_t = \mu + \beta t + Y_t^* \quad \text{for} \quad t = 0,\ldots,T,\\
&\qquad\qquad \text{for some} \quad \mu,\beta \in R \quad \text{and} \quad \alpha \in (-1,1].
\end{aligned}
$$

Models 1–3 can be written equivalently as:

$$
\begin{aligned}
(2.3)\quad &\text{Model 1:}\quad Y_t = \alpha Y_{t-1} + U_t \quad \text{for} \quad t = 1,\ldots,T \quad \text{and} \quad \alpha \in (-1,1),\\
&\text{Model 2:}\quad Y_t = \tilde{\mu} + \alpha Y_{t-1} + U_t \quad \text{for} \quad t = 1,\ldots,T,\\
&\qquad\qquad \text{where} \quad \tilde{\mu} = \mu(1-\alpha) \quad \text{and} \quad \alpha \in (-1,1],\\
&\text{Model 3:}\quad Y_t = \tilde{\mu} + \tilde{\beta} t + \alpha Y_{t-1} + U_t \quad \text{for} \quad t = 1,\ldots,T,\\
&\qquad \text{where} \quad \tilde{\mu} = \mu(1-\alpha) + \alpha\beta, \quad \tilde{\beta} = \beta(1-\alpha), \quad \text{and} \quad \alpha \in (-1,1].
\end{aligned}
$$

In (2.3), $\{U_t\}$ is as in (2.1), $Y_0 \sim N(0, \sigma^2/(1-\alpha^2))$ in Model 1, $Y_0 \sim N(\mu, \sigma^2/(1-\alpha^2))$ in Models 2 and 3 if $\alpha \in (-1,1)$, and $Y_0$ is an arbitrary constant or random variable in Models 2 and 3 if $\alpha = 1$.[3]

Below we use the term *sample size* to denote the number of $Y_t$ observations, viz., $T+1$.

In Model 1, $\{Y_t\colon t = 1,\ldots,T\}$ is a strictly stationary, normal, AR(1) process with mean zero. In this model, the nonstationary parameter value $\alpha = 1$ is excluded. The reason is that the LS estimator of $\alpha$ has distribution that depends on the initial condition $Y_0$ in this case and in consequence the exact procedures that are introduced below do not apply when $\alpha = 1$. Since Model 1 is not of great practical importance—one rarely assumes the mean is known and equals zero—this is not a serious restriction. Model 1 is included in the discussion for comparative purposes and for completeness, rather than for its applicability in practice.

In Model 2, if $\alpha \in (-1,1)$, $\{Y_t\colon t = 1,\ldots,T\}$ is a strictly stationary, normal, AR(1) process with mean $\mu$. In this model, if $\alpha = 1$, $\{Y_t\colon t = 1,\ldots,T\}$ is a normal random walk with arbitrary initial condition. In Model 3, if $\alpha \in (-1,1)$, $\{Y_t\colon t = 1,\ldots,T\}$ is a strictly stationary, normal, AR(1) process around a deterministic trend line with intercept $\mu$ and slope $\beta$. In this model, if $\alpha = 1$, $\{Y_t\colon t = 1,\ldots,T\}$ is a normal random walk with drift $\beta$ and arbitrary initial condition $Y_0$. The results of this paper also apply to models with higher-order time trends.

[3] Model 3 is quite similar to that considered by Bhargava (1986) who is concerned with tests of a unit root. Model 3 is less restrictive, however, in that it allows the initial observation $Y_0$ to be arbitrary when $\alpha = 1$, whereas Bhargava requires it to have a particular normal distribution. On the other hand, Bhargava considers cases where $\alpha > 1$. The approach of the present paper could also be extended to cover cases where $\alpha > 1$, but only by imposing a very specific assumption on the initial observation $Y_0$ (as in Bhargava (1986)).

For brevity, we only consider Models 1–3 here. The extension of the results is straightforward.

Let $\hat{\alpha}_{LSj}$ denote the LS estimator of $\alpha$ corresponding to Model $j$ given in (2.3). That is, $\hat{\alpha}_{LS1}, \hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ are the LS estimators of $\alpha$ from the regression of $Y_t$ on $Y_{t-1}, (1, Y_{t-1})$, and $(1, t, Y_{t-1})$, respectively, for $t = 1, \ldots, T$. We note the following: For $j = 1, 2, 3$:

PROPERTY OF $\hat{\alpha}_{LSj}$: *The distribution of $\hat{\alpha}_{LSj}$ depends only on $\alpha$ when Model $j$ is correct. In particular, it does not depend on $\sigma^2$ in Model 1, $(\sigma^2, \mu)$ in Model 2, or $(\sigma^2, \mu, \beta)$ in Model 3, nor on the value of $Y_0$ when $\alpha = 1$ in Models 2 and 3.*

A proof of this property is given in Appendix A. Similar distributional invariance results for various estimators and test statistics in models similar to Models 1, 2, and 3 have been pointed out by Dickey and Fuller (1979), DeJong et al. (1992a), and Kiviet and Phillips (1990), among others. The above property of $\hat{\alpha}_{LSj}$ is exploited in the construction of median-unbiased estimators of $\alpha$ in Section 3 below.

## 3. MEDIAN-UNBIASED ESTIMATION OF THE AUTOREGRESSIVE PARAMETER

This section introduces median-unbiased estimators of $\alpha$ for Models 1–3.

A number $m$ is a *median* of a random variable $X$ if

$$(3.1) \qquad P(X \geq m) \geq 1/2 \quad \text{and} \quad P(X \leq m) \geq 1/2.$$

This definition of a median allows for nonuniqueness, but all of the medians considered here will be unique. The definition also allows for the median of $X$ to be a probability mass point of $X$. This feature of the definition will be used here. If a median $m$ of $X$ is not a probability mass point, then $P(X > m) = P(X < m) = 1/2$.

Let $\hat{\alpha}$ be an estimator of the parameter $\alpha$. By definition, $\hat{\alpha}$ is *median-unbiased* for $\alpha$ if the true parameter $\alpha$ is a median of $\hat{\alpha}$ for each $\alpha$ in the parameter space. The condition of median-unbiasedness has the intuitive impartiality property that the probability of underestimation equals the probability of overestimation. This holds unless the true parameter value is estimated with positive probability and in this case the probabilities of underestimation and overestimation are each less than one half. When the magnitude of a parameter is a contentious issue, such as in the (trend) stationary versus unit root debate, this impartiality property is quite useful.

Median-unbiasedness is a special case of the concept of *risk-unbiasedness* when loss is given by absolute error. In particular, $\hat{\alpha}$ is median-unbiased for $\alpha$, if and only if it has the property that

$$(3.2) \qquad E_\alpha|\hat{\alpha} - \alpha| \leq E_\alpha|\hat{\alpha} - \alpha'| \quad \text{for all} \quad \alpha, \alpha' \in A,$$

where $A$ denotes the parameter space of $\alpha$ and $E_\alpha$ denotes the expectation operator when $\alpha$ is the true parameter value; see Lehmann (1959, p. 22). In

words, $\hat{\alpha}$ is median-unbiased if and only if the distance between $\hat{\alpha}$ and the true parameter on average is less than or equal to the distance between $\hat{\alpha}$ and any other parameter value. In this sense, the value that $\hat{\alpha}$ is best at estimating is the true value $\alpha$ regardless of what $\alpha$ is.

The condition of median-unbiasedness is often more useful than that of mean-unbiasedness when the parameter space is bounded or when the distributions of estimators are skewed and/or kurtotic. When the parameter space is bounded and closed and estimators take values in the parameter space, it is impossible to have a mean-unbiased estimator because all estimators are biased at extreme boundary points. Boundary points do not present problems, however, for the condition of median-unbiasedness: If an estimator is median-unbiased for a parameter space $A \subset R$, then the estimator restricted to a closed subset $A^*$ of $A$ is median-unbiased for the restricted parameter space $A^*$. (The method of restricting the estimator, say $\hat{\alpha}$, to $A^*$ is to set $\hat{\alpha}$ equal to the nearest element of $A^*$ that is larger or smaller than $\hat{\alpha}$.) Next, when estimators have asymmetric distributions, there is no unambiguous measure of the centers of their distributions. In this case, the median may be a preferred measure to the mean, especially in case of kurtosis, because the median is less sensitive to the tails of the distribution.

We note that in the classical normal linear regression model with fixed regressors the LS estimator is median-unbiased. In fact, it is the best median-unbiased estimator for a wide variety of loss functions (see Andrews and Phillips (1987)). In the AR/UR model, on the other hand, the LS estimator is not median-unbiased, and hence, does not possess the same optimality properties.

We now discuss an exact method for median-bias correcting an estimator. (This method is not original to the present paper; e.g., it more or less corresponds to the method discussed by Lehmann (1959, Sec. 3.5, p. 83), but the application of the method below to the LS estimator in Models 1–3 is original.) Suppose $\hat{\alpha}$ is an estimator whose median function $m(\alpha)$ $(= m_T(\alpha))$ is uniquely defined and is strictly increasing on the parameter space $A$ which is a finite interval, say $(-1, 1]$. (By definition, $m(\alpha)$ is the unique median of $\hat{\alpha}$ when $\alpha$ is the true parameter $\forall \alpha \in (-1, 1]$.) Then $\hat{\alpha}_U$ is a median-unbiased estimator of $\alpha$, where

$$\hat{\alpha}_U = \begin{cases} 1 & \text{if } \hat{\alpha} > m(1), \\ m^{-1}(\hat{\alpha}) & \text{if } m(-1) < \hat{\alpha} \leq m(1), \\ -1 & \text{if } \hat{\alpha} \leq m(-1), \end{cases} \tag{3.3}$$

$m(-1) = \lim_{\alpha \to -1} m(\alpha)$, and $m^{-1}$: $(m(-1), m(1)] \to (-1, 1]$ is the inverse function of $m(\cdot)$ that satisfies $m^{-1}(m(\alpha)) = \alpha$ for $\alpha \in (-1, 1]$. (Since the parameter space is $(-1, 1)$ in Model 1 rather than $(-1, 1]$, $m(1)$ needs to be defined as $\lim_{\alpha \to 1} m(\alpha)$ in this case.)

To show that $\hat{\alpha}_U$ is median-unbiased, we note that $\forall \alpha \in (-1, 1]$ $\hat{\alpha}_U \geq \alpha$ iff $m(\hat{\alpha}_U) \geq m(\alpha)$ iff $\hat{\alpha} \geq m(\alpha)$, where the first "iff" holds because $m(\cdot)$ is increas-

ing and the second holds by the definition of $\hat{\alpha}_U$ in (3.3). Since the same result holds with $\geq$ replaced by $\leq$ throughout, median-unbiasedness follows.

The above method of bias correction can be applied to the estimators $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ defined in Section 2. Each of these estimators has a distribution that depends only on $\alpha$ (and the sample size $T+1$) and numerical evaluation of their median functions $m(\alpha)$ show the latter to be strictly increasing, as one would expect.[4] In consequence, we can define three corresponding median-unbiased estimators $\hat{\alpha}_{U1}$, $\hat{\alpha}_{U2}$, and $\hat{\alpha}_{U3}$ using the formula given in (3.3) with $\hat{\alpha}$ replaced by $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ respectively. These estimators can be computed given $\hat{\alpha}_{LSj}$ if the function $m^{-1}(\cdot)$, or equivalently, $m(\cdot)$ is known.

Tables I, II, and III provide values of the median function $m(\alpha)$ of $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ for a grid of $\alpha$ values and various sample sizes. The numerical method used to construct the tables is described in Appendix B.

To illustrate the use of the tables, consider the estimator $\hat{\alpha}_{LS3}$ for Model 3 and sample size $T+1=60$. As shown in Table III, $m(1)=.853$, so any value of $\hat{\alpha}_{LS3}$ that is $\geq .853$ corresponds to $\hat{\alpha}_{U3}=1$. Similarly, $m(-1)=-.997$, so any value of $\hat{\alpha}_{LS3}$ that is $\leq -.997$ corresponds to $\hat{\alpha}_{U3}=-1$. For any value of $\hat{\alpha}_{LS3}$ between $-.997$ and .853, one finds the entry in the $m(\alpha)$ column (i.e., the .5 quantile column) that equals $\hat{\alpha}_{LS3}$ and the $\alpha$ value that corresponds to this entry is $\hat{\alpha}_{U3}$. That is, $\hat{\alpha}_{U3}$ is chosen such that $m(\hat{\alpha}_{U3})=\hat{\alpha}_{LS3}$. For example, if $\hat{\alpha}_{LS3}=.80$, then $\hat{\alpha}_{U3}=.90$. Since the grid of $\alpha$ values given in the tables is finite, interpolation between $\alpha$ values may be needed. Based on a more complete set of tables, we find that for $\alpha$ values of $-.2$ or greater, linear interpolation incurs an error of .001 or less. Similarly, for sample sizes not given in the tables, interpolation is required. Alternatively, $m(\alpha)$ can be calculated numerically or by simulation for a particular sample size of interest.

Although we do not pursue the idea here, we note that one also could calculate a median-unbiased estimator of $\alpha$ based on the $t$ statistic for $\alpha$, rather than on $\hat{\alpha}_{LSj}$, using the method described above. We also note that the method considered here can be applied to the model $Y_t=X_t'\beta+Y_t^*$, $Y_t^*=\alpha Y_{t-1}^*+U_t$ where $X_t$ is a fixed regressor vector. One takes $\hat{\alpha}$ to be the LS estimator from the regression of $\tilde{Y}_t$ on $\tilde{Y}_{t-1}$, where $\tilde{Y}_t$ is the residual from the regression of $Y_t$ on $X_t$. The method can be applied to the model $Y_t=X_t'\beta+\alpha Y_{t-1}+U_t$ if one takes $\hat{\alpha}$ to be the LS estimator from the regression of $Y_t$ on $X_t$, $Y_{t-1}$, and an additional regressor vector $Z_t$ that depends on $\{X_1,\ldots,X_T\}$ and is designed to leave the distribution of $\hat{\alpha}$ independent of $\beta$. See Kiviet and Phillips (1990) for the specification of $Z_t$.

[4] It would be useful to have an analytic proof that the median of $\hat{\alpha}_{LSj}$ is strictly increasing for $\alpha\in(-1,1]$ for all $T\geq 4$ and $j=1,2,3$, but we do not have such a proof at present. We do know, however, that the more general proposition that all of the quantiles of $\hat{\alpha}_{LSj}$ are strictly increasing for $\alpha\in(-1,1]$ is false. Numerical calculations of the .95 quantile function of $\hat{\alpha}_{LS1}$ for sample sizes $T+1=10,\ldots,60$ reveal that the function dips very slightly for $\alpha$ very close to one; see Table 1. The numerical calculations have been substantiated by simulation.

TABLE I

QUANTILES OF THE LEAST SQUARES ESTIMATOR OF $\alpha$ FOR MODEL 1[a]

| | $T+1=60$ | | | $T+1=100$ | | |
|---|---|---|---|---|---|---|
| $\alpha \setminus$ Quantile | .05 | .5 | .95 | .05 | .5 | .95 |
| .00 | −.211 | .000 | .211 | −.164 | .000 | .164 |
| .05 | −.164 | .049 | .258 | −.115 | .050 | .212 |
| .10 | −.116 | .098 | .304 | −.067 | .099 | .259 |
| .15 | −.068 | .148 | .350 | −.017 | .149 | .307 |
| .20 | −.019 | .197 | .395 | .032 | .198 | .353 |
| .25 | .030 | .246 | .440 | .082 | .248 | .400 |
| .30 | .080 | .295 | .484 | .132 | .297 | .446 |
| .35 | .130 | .344 | .528 | .183 | .347 | .491 |
| .40 | .181 | .393 | .571 | .235 | .396 | .536 |
| .45 | .233 | .443 | .614 | .286 | .446 | .581 |
| .50 | .285 | .492 | .656 | .339 | .495 | .625 |
| .55 | .338 | .541 | .697 | .392 | .545 | .669 |
| .60 | .392 | .590 | .738 | .445 | .594 | .711 |
| .65 | .446 | .640 | .777 | .500 | .644 | .754 |
| .70 | .502 | .689 | .816 | .555 | .693 | .795 |
| .75 | .559 | .738 | .854 | .611 | .743 | .835 |
| .80 | .618 | .787 | .890 | .669 | .792 | .874 |
| .85 | .678 | .837 | .924 | .728 | .842 | .912 |
| .90 | .719 | .886 | .957 | .790 | .891 | .948 |
| .93 | .784 | .917 | .975 | .829 | .921 | .968 |
| .95 | .813 | .937 | .987 | .844 | .941 | .980 |
| .97 | .846 | .958 | .998 | .886 | .962 | .992 |
| .99 | .891 | .981 | 1.009 | .920 | .983 | 1.003 |
| .995 | .910 | .989 | 1.011 | .935 | .990 | 1.006 |
| .999 | .944 | .997 | 1.010 | .955 | .997 | 1.007 |

[a] A symmetry argument similar to that of Fuller (1976, pp. 370–372) shows the distribution of $\hat{\alpha}_{LS1}$ for $\alpha$ negative to be a mirror image about 0 of its distribution for $\alpha$ positive. Thus, the .05, .5, and .95 quantiles of $\hat{\alpha}_{LS1}$ for $\alpha \in (-.999, 0)$ equal the .95, .5, and .05 quantiles, respectively, of $\hat{\alpha}_{LS1}$ for $-\alpha \in (0, .999)$. In consequence, Table I provides the quantiles of $\hat{\alpha}_{LS1}$ for all values of $\alpha \in (-.999, .999)$.

## 4. EXACT CONFIDENCE INTERVALS FOR THE AUTOREGRESSIVE PARAMETER

This section derives exact confidence intervals for the parameter $\alpha$ in Models 1–3. These confidence intervals can be used by themselves, in conjunction with the median-unbiased estimator $\hat{\alpha}_{Uj}$ of $\alpha$ (to provide a measure of accuracy of $\hat{\alpha}_{Uj}$), or to construct exact one- or two-sided tests of $H_0$: $\alpha = \alpha_0$ for arbitrary $\alpha_0 \in (-1, 1]$.

For model $j$ we construct the confidence interval using $\hat{\alpha}_{LSj}$ for $j = 1, 2, 3$. As shown in Section 2, the distribution of $\hat{\alpha}_{LSj}$ depends only on the parameter $\alpha$ for which we wish to construct the confidence interval. In addition, its distribution is absolutely continuous and has support $R$ for all $\alpha$ in $(-1, 1]$ ($(-1, 1)$ for Model 1). Let $q_p(\alpha)$ denote the $p$th quantile function of $\hat{\alpha}_{LSj}$. By definition, for fixed $p \in (0, 1)$, $P_\alpha(\hat{\alpha}_{LSj} \leq q_p(\alpha)) = p$, where $P_\alpha$ denotes the probability distribution when $\alpha$ is the true value.

A $100(1-p)\%$ confidence interval (set) for $\alpha$ in Model $j$ is given by the realization of the set

$$\{\alpha \in [-1, 1]: q_{p_1}(\alpha) \leq \hat{\alpha}_{LSj} \leq q_{p_2}(\alpha)\}, \tag{4.1}$$

TABLE II

QUANTILES OF THE LEAST SQUARES ESTIMATOR OF $\alpha$ FOR MODEL 2

| | $T+1=40$ | | | $T+1=50$ | | | $T+1=60$ | | | $T+1=70$ | | | $T+1=80$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$/Quantile | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 |
| −.999 | −1.014 | −.997 | −.928 | −1.012 | −.997 | −.938 | −1.010 | −.997 | −.945 | −1.009 | −.997 | −.950 | −1.008 | −.997 | −.954 |
| −.80 | −.907 | −.787 | −.572 | −.898 | −.789 | −.602 | −.892 | −.791 | −.624 | −.887 | −.792 | −.640 | −.882 | −.793 | −.653 |
| −.60 | −.769 | −.596 | −.349 | −.754 | −.597 | −.379 | −.743 | −.597 | −.401 | −.734 | −.598 | −.417 | −.726 | −.598 | −.431 |
| −.40 | −.616 | −.406 | −.145 | −.595 | −.405 | −.173 | −.580 | −.404 | −.193 | −.567 | −.403 | −.209 | −.557 | −.403 | −.222 |
| −.20 | −.454 | −.216 | .048 | −.427 | −.213 | .023 | −.407 | −.210 | .004 | −.392 | −.209 | −.010 | −.380 | −.208 | −.022 |
| .00 | −.282 | −.026 | .232 | −.251 | −.021 | .211 | −.228 | −.017 | .194 | −.210 | −.015 | .181 | −.195 | −.013 | .170 |
| .10 | −.194 | .069 | .321 | −.160 | .075 | .302 | −.135 | .079 | .286 | −.116 | .082 | .274 | −.101 | .085 | .264 |
| .20 | −.103 | .164 | .408 | −.067 | .171 | .390 | −.041 | .176 | .377 | −.020 | .180 | .366 | −.005 | .182 | .357 |
| .30 | −.011 | .258 | .493 | .028 | .267 | .477 | .056 | .273 | .465 | .077 | .277 | .455 | .093 | .280 | .447 |
| .40 | .083 | .353 | .575 | .124 | .363 | .562 | .154 | .369 | .551 | .176 | .374 | .543 | .193 | .377 | .536 |
| .50 | .180 | .447 | .655 | .223 | .458 | .644 | .254 | .465 | .635 | .276 | .470 | .628 | .294 | .474 | .622 |
| .60 | .278 | .541 | .732 | .324 | .554 | .724 | .356 | .562 | .717 | .379 | .567 | .711 | .398 | .572 | .706 |
| .70 | .378 | .635 | .806 | .427 | .649 | .800 | .460 | .658 | .795 | .485 | .664 | .791 | .504 | .669 | .788 |
| .80 | .480 | .727 | .875 | .532 | .743 | .872 | .567 | .753 | .870 | .593 | .760 | .867 | .613 | .765 | .865 |
| .85 | .532 | .772 | .908 | .585 | .789 | .906 | .622 | .800 | .905 | .648 | .807 | .903 | .668 | .813 | .902 |
| .90 | .582 | .816 | .939 | .638 | .834 | .938 | .676 | .846 | .938 | .704 | .854 | .937 | .724 | .861 | .937 |
| .93 | .612 | .841 | .958 | .669 | .860 | .957 | .709 | .873 | .957 | .737 | .882 | .957 | .758 | .888 | .956 |
| .97 | .649 | .872 | .983 | .706 | .893 | .982 | .750 | .906 | .982 | .780 | .916 | .982 | .802 | .923 | .982 |
| .99 | .666 | .886 | .994 | .727 | .907 | .994 | .769 | .921 | .994 | .799 | .931 | .994 | .822 | .939 | .994 |
| 1.0 | .674 | .893 | .999 | .735 | .914 | .999 | .777 | .928 | .999 | .807 | .938 | .999 | .831 | .946 | .999 |

| | $T+1=90$ | | | $T+1=100$ | | | $T+1=125$ | | | $T+1=150$ | | | $T+1=200$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$/Quantile | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 |
| −.999 | −1.007 | −.997 | −.958 | −1.006 | −.997 | −.961 | −1.005 | −.997 | −.966 | −1.004 | −.997 | −.970 | −1.003 | −.998 | −.976 |
| −.80 | −.879 | −.794 | −.663 | −.876 | −.794 | −.672 | −.869 | −.795 | −.688 | −.865 | −.796 | −.700 | −.858 | −.797 | −.715 |
| −.60 | −.720 | −.598 | −.441 | −.715 | −.598 | −.450 | −.704 | −.599 | −.468 | −.696 | −.599 | −.480 | −.684 | −.599 | −.498 |
| −.40 | −.549 | −.402 | −.233 | −.542 | −.402 | −.242 | −.528 | −.402 | −.259 | −.517 | −.401 | −.272 | −.502 | −.401 | −.289 |
| −.20 | −.370 | −.207 | −.032 | −.361 | −.206 | −.041 | −.344 | −.205 | −.057 | −.332 | −.204 | −.069 | −.314 | −.203 | −.087 |
| .00 | −.184 | −.011 | .161 | −.174 | −.010 | .154 | −.155 | −.008 | .139 | −.141 | −.007 | .127 | −.121 | −.005 | .111 |
| .10 | −.088 | .086 | .256 | −.078 | .088 | .249 | −.058 | .090 | .234 | −.043 | .092 | .224 | −.023 | .094 | .208 |
| .20 | .008 | .184 | .349 | .019 | .186 | .342 | .040 | .189 | .329 | .055 | .191 | .319 | .076 | .193 | .304 |
| .30 | .107 | .282 | .440 | .118 | .284 | .434 | .140 | .287 | .422 | .155 | .289 | .413 | .177 | .292 | .399 |
| .40 | .207 | .380 | .530 | .219 | .382 | .524 | .241 | .385 | .513 | .257 | .388 | .505 | .278 | .391 | .493 |
| .50 | .309 | .477 | .617 | .321 | .480 | .612 | .343 | .484 | .603 | .359 | .486 | .596 | .381 | .490 | .585 |
| .60 | .413 | .575 | .702 | .425 | .577 | .699 | .448 | .582 | .691 | .464 | .585 | .685 | .486 | .589 | .676 |
| .70 | .519 | .672 | .785 | .532 | .675 | .782 | .555 | .680 | .776 | .571 | .684 | .771 | .593 | .688 | .764 |
| .80 | .628 | .769 | .863 | .641 | .773 | .862 | .664 | .778 | .858 | .681 | .782 | .855 | .702 | .787 | .850 |
| .85 | .684 | .818 | .901 | .697 | .821 | .900 | .721 | .827 | .897 | .737 | .831 | .895 | .758 | .836 | .891 |
| .90 | .741 | .865 | .936 | .754 | .869 | .936 | .778 | .876 | .934 | .794 | .880 | .933 | .815 | .885 | .931 |
| .93 | .775 | .893 | .956 | .788 | .897 | .956 | .813 | .904 | .955 | .829 | .909 | .955 | .850 | .915 | .953 |
| .97 | .820 | .929 | .981 | .834 | .933 | .981 | .859 | .941 | .981 | .876 | .947 | .981 | .898 | .953 | .981 |
| .99 | .840 | .945 | .994 | .854 | .950 | .994 | .881 | .958 | .994 | .898 | .964 | .994 | .921 | .971 | .994 |
| 1.0 | .849 | .952 | .999 | .863 | .957 | .999 | .890 | .965 | .999 | .908 | .971 | .999 | .931 | .978 | .999 |

TABLE III

QUANTILES OF THE LEAST SQUARES ESTIMATOR OF $\alpha$ FOR MODEL 3

| | $T+1=40$ | | | $T+1=50$ | | | $T+1=60$ | | | $T+1=70$ | | | $T+1=80$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$/Quantile | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 |
| −.999 | −1.015 | −.997 | −.930 | −1.012 | −.997 | −.939 | −1.010 | −.997 | −.945 | −1.009 | −.997 | −.951 | −1.008 | −.997 | −.955 |
| −.80 | −.910 | −.793 | −.581 | −.901 | −.794 | −.609 | −.894 | −.794 | −.629 | −.889 | −.795 | −.645 | −.884 | −.796 | −.657 |
| −.60 | −.776 | −.607 | −.364 | −.760 | −.605 | −.390 | −.748 | −.604 | −.410 | −.738 | −.603 | −.425 | −.730 | −.603 | −.437 |
| −.40 | −.628 | −.421 | −.163 | −.605 | −.417 | −.187 | −.588 | −.414 | −.205 | −.575 | −.412 | −.219 | −.564 | −.410 | −.231 |
| −.20 | −.471 | −.237 | .026 | −.441 | −.229 | .006 | −.420 | −.224 | −.010 | −.403 | −.220 | −.023 | −.389 | −.218 | −.033 |
| .00 | −.307 | −.052 | .207 | −.270 | −.041 | .191 | −.244 | −.034 | .177 | −.224 | −.029 | .167 | −.208 | −.026 | .158 |
| .10 | −.222 | .040 | .295 | −.182 | .052 | .280 | −.154 | .061 | .269 | −.132 | .066 | .259 | −.115 | .071 | .251 |
| .20 | −.136 | .132 | .380 | −.093 | .146 | .368 | −.062 | .155 | .358 | −.039 | .162 | .349 | −.021 | .167 | .342 |
| .30 | −.048 | .223 | .463 | −.002 | .239 | .454 | .031 | .250 | .445 | .056 | .257 | .438 | .075 | .263 | .432 |
| .40 | .040 | .314 | .544 | .090 | .332 | .537 | .126 | .344 | .531 | .152 | .352 | .525 | .172 | .359 | .520 |
| .50 | .130 | .405 | .623 | .184 | .425 | .619 | .222 | .438 | .614 | .249 | .447 | .610 | .271 | .454 | .606 |
| .60 | .220 | .494 | .699 | .279 | .517 | .697 | .319 | .532 | .695 | .349 | .542 | .692 | .371 | .550 | .690 |
| .70 | .310 | .582 | .771 | .374 | .608 | .773 | .417 | .624 | .773 | .449 | .636 | .772 | .473 | .644 | .771 |
| .80 | .397 | .665 | .838 | .467 | .695 | .843 | .515 | .715 | .846 | .549 | .728 | .847 | .575 | .738 | .848 |
| .85 | .438 | .704 | .869 | .512 | .737 | .876 | .562 | .758 | .880 | .599 | .773 | .883 | .626 | .784 | .884 |
| .90 | .475 | .740 | .898 | .554 | .776 | .907 | .607 | .799 | .912 | .646 | .815 | .916 | .675 | .827 | .918 |
| .93 | .494 | .758 | .913 | .576 | .796 | .923 | .632 | .821 | .929 | .672 | .839 | .934 | .703 | .852 | .937 |
| .97 | .513 | .776 | .930 | .599 | .817 | .941 | .658 | .845 | .949 | .701 | .864 | .954 | .734 | .879 | .958 |
| .99 | .518 | .781 | .935 | .605 | .823 | .947 | .665 | .851 | .955 | .709 | .872 | .961 | .743 | .887 | .966 |
| 1.0 | .519 | .782 | .935 | .606 | .824 | .948 | .666 | .853 | .956 | .711 | .873 | .962 | .745 | .889 | .967 |

| $\alpha$/Quantile | $T+1=90$ | | | $T+1=100$ | | | $T+1=125$ | | | $T+1=150$ | | | $T+1=200$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 | .05 | .5 | .95 |
| −.999 | −1.007 | −.997 | −.958 | −1.006 | −.997 | −.961 | −1.005 | −.997 | −.966 | −1.004 | −.997 | −.970 | −1.003 | −.998 | −.976 |
| −.80 | −.880 | −.796 | −.667 | −.877 | −.796 | −.675 | −.871 | −.797 | −.690 | −.866 | −.798 | −.702 | −.858 | −.798 | −.717 |
| −.60 | −.723 | −.603 | −.447 | −.718 | −.602 | −.456 | −.707 | −.602 | −.472 | −.698 | −.601 | −.484 | −.686 | −.601 | −.500 |
| −.40 | −.555 | −.409 | −.240 | −.547 | −.408 | −.248 | −.532 | −.407 | −.264 | −.521 | −.405 | −.276 | −.505 | −.404 | −.293 |
| −.20 | −.378 | −.216 | −.042 | −.368 | −.214 | −.049 | −.350 | −.211 | −.064 | −.337 | −.209 | −.075 | −.318 | −.207 | −.091 |
| .00 | −.195 | −.023 | .150 | −.184 | −.020 | .144 | −.163 | −.016 | .131 | −.147 | −.013 | .121 | −.126 | −.010 | .106 |
| .10 | −.101 | .074 | .244 | −.089 | .077 | .238 | −.067 | .081 | .226 | −.051 | .084 | .217 | −.029 | .088 | .203 |
| .20 | −.006 | .170 | .336 | .007 | .173 | .331 | .030 | .179 | .320 | .047 | .182 | .311 | .070 | .187 | .299 |
| .30 | .091 | .267 | .427 | .104 | .270 | .422 | .128 | .276 | .412 | .146 | .280 | .405 | .170 | .285 | .393 |
| .40 | .189 | .363 | .516 | .202 | .367 | .512 | .228 | .374 | .503 | .246 | .378 | .497 | .271 | .384 | .487 |
| .50 | .288 | .460 | .603 | .302 | .464 | .599 | .329 | .471 | .593 | .348 | .476 | .587 | .373 | .482 | .579 |
| .60 | .389 | .556 | .687 | .404 | .560 | .685 | .432 | .568 | .680 | .451 | .574 | .676 | .476 | .581 | .669 |
| .70 | .492 | .651 | .769 | .508 | .656 | .768 | .536 | .665 | .765 | .556 | .671 | .762 | .582 | .679 | .757 |
| .80 | .596 | .745 | .848 | .612 | .751 | .848 | .643 | .762 | .847 | .663 | .769 | .845 | .689 | .777 | .843 |
| .85 | .648 | .792 | .885 | .665 | .798 | .885 | .696 | .809 | .886 | .717 | .817 | .885 | .744 | .826 | .884 |
| .90 | .698 | .836 | .920 | .716 | .844 | .921 | .749 | .856 | .923 | .771 | .864 | .923 | .799 | .874 | .923 |
| .93 | .727 | .861 | .939 | .746 | .869 | .941 | .780 | .883 | .943 | .803 | .892 | .945 | .832 | .902 | .946 |
| .97 | .759 | .890 | .961 | .780 | .899 | .964 | .818 | .915 | .968 | .843 | .925 | .970 | .874 | .938 | .973 |
| .99 | .770 | .899 | .969 | .791 | .909 | .972 | .831 | .926 | .977 | .857 | .938 | .980 | .891 | .952 | .984 |
| 1.0 | .772 | .901 | .971 | .793 | .911 | .974 | .833 | .928 | .979 | .860 | .940 | .982 | .894 | .955 | .987 |

where $p_1 > 0$, $p_2 > 0$, and $p = p_1 + 1 - p_2$. This set has the correct coverage probability because

$$(4.2) \quad P_{\alpha_0}\left(\alpha_0 \in \left\{\alpha \in [-1,1]: q_{p_1}(\alpha) \le \hat{\alpha}_{LSj} \le q_{p_2}(\alpha)\right\}\right)$$
$$= P_{\alpha_0}\left(q_{p_1}(\alpha_0) \le \hat{\alpha}_{LSj} \le q_{p_2}(\alpha_0)\right) = 1 - p$$

for all $\alpha_0 \in (-1,1]$. (Note that the method used here of constructing confidence intervals is time honored, only the application of it in this context is original.)

If the quantile functions $q_{p_1}(\alpha)$ and $q_{p_2}(\alpha)$ are strictly increasing in $\alpha$ (as they are in all the cases tabulated in Tables II and III), then the set in (4.1) equals the interval $\{\alpha: \hat{c}_L \le \alpha \le \hat{c}_U\}$, where $\hat{c}_L$ and $\hat{c}_U$ are defined using the same formula (3.3) as for $\hat{\alpha}_U$, but with $\hat{\alpha}$, $m(\cdot)$, and $m^{-1}(\cdot)$ replaced by $\hat{\alpha}_{LSj}$, $q_{p_2}(\cdot)$, and $q_{p_2}^{-1}(\cdot)$ for $\hat{c}_L$, with $\hat{c}_L > 1$ rather than $\hat{c}_L = 1$ if $\hat{\alpha}_{LSj} > q_{p_2}(1)$, and with $\hat{\alpha}$, $m(\cdot)$, and $m^{-1}(\cdot)$ replaced by $\hat{\alpha}_{LSj}$, $q_{p_1}(\cdot)$, and $q_{p_1}^{-1}(\cdot)$ for $\hat{c}_U$.[5] By definition, for $i = 1,2$, $q_{p_i}(-1) = \lim_{\alpha \to -1} q_{p_i}(\alpha)$ and $q_{p_i}^{-1}(\cdot)$: $(q_{p_i}(-1), q_{p_i}(1)] \to (-1,1]$ is the inverse function of $q_{p_i}(\cdot)$ that satisfies $q_{p_i}^{-1}(q_{p_i}(\alpha)) = \alpha$ for $\alpha \in (-1,1]$. (For Model 1, $q_{p_i}(1) = \lim_{\alpha \to 1} q_{p_i}(\alpha)$ and the domain of $q_{p_i}^{-1}$ is $(q_{p_i}(-1), q_{p_i}(1))$.) Note that $\hat{\alpha}_U$ is just the lower and upper bounds of the two one-sided level .5 confidence intervals for $\alpha$ when $m(\cdot)$ is strictly increasing (provided one defines $\hat{\alpha}_U = 1$ if $\hat{c}_L > 1$).

Tables I–III provide values of the quantile functions $q_{.05}(\alpha)$ and $q_{.95}(\alpha)$ of $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ for different sample sizes. The values can be used to construct two-sided 90% confidence intervals and one-sided 95% confidence intervals for $\alpha$. (For a lower bound one-sided confidence interval, one takes $p_1 = 0$ and the interval is of the form $[\hat{c}_L, 1]$. Analogously, for an upper bound one-sided confidence interval, one takes $p_2 = 1$ and the interval is of the form $(-1, \hat{c}_U]$.)

The interval endpoints $\hat{c}_L$ and $\hat{c}_U$ are obtained from Tables I–III in exactly the same way as is the median-unbiased estimator $\hat{\alpha}_L$. For example, consider a 90% two-sided confidence interval for $\alpha$ in Model 3 with sample size 60. As shown in Table III, $q_{.95}(1) = .956$, so any value of $\hat{\alpha}_{LS3}$ that is $\ge .956$ corresponds to $\hat{c}_L > 1$. Similarly, $q_{.95}(-1) = -.945$, so any value of $\hat{\alpha}_{LS3}$ that is $\le -.945$ corresponds to $\hat{c}_L = -1$. For any value of $\hat{\alpha}_{LS3}$ between $-.945$ and .956, one finds the entry in the $q_{.95}(\alpha)$ column (i.e., the .95 quantile column),

[5] As defined, if $\hat{\alpha}_{LSj} > q_{p_2}(1)$, then $\hat{c}_L > 1$, $\hat{c}_U = 1$, and the confidence interval $[\hat{c}_L, \hat{c}_U]$ equals the null set. This definition guarantees that the coverage probability of the random interval $[\hat{c}_L, \hat{c}_U]$ is exactly $p$ when $\alpha = 1$ just as it is for all other values of $\alpha$. An alternative, and perhaps more natural, definition of $\hat{c}_L$ is to set it equal to 1 when $\hat{\alpha}_{LSj} > q_{p_2}(1)$ and then the confidence interval is the single point {1} in this case. If this is done, the coverage probability of the random interval $[\hat{c}_L, \hat{c}_U]$ is exactly $1 - p$ for all $\alpha \in (-1,1)$, but is $1 - p_1 (> 1 - p)$ for $\alpha = 1$ (since there is no way for the confidence interval to miss the point $\alpha = 1$ by being to the right of 1). This asymmetry between the coverage probabilities of the confidence interval for the (trend) stationary and nonstationary cases may be deemed undesirable if attention is focussed on which of the two models is correct. If it *is* deemed to be undesirable, then one can define the confidence interval as in the text and view an occurrence of a null confidence set as evidence against the model restriction that $\alpha \in (-1,1]$, since the test "reject $H_0$: $\alpha \in (-1,1]$ in favor of $H_1$: $\alpha > 1$ if $\hat{\alpha}_{LSj} > q_{p_2}(1)$" is a valid level $p_2$ significance test.

that equals $\hat{\alpha}_{LS3}$ and the $\alpha$ value that corresponds to this entry is $\hat{c}_L$. For instance, if $\hat{\alpha}_{LS3} = .80$, then $\hat{c}_L = .74$ (utilizing a linear interpolation). One obtains $\hat{c}_U$ in exactly the same manner using the column of $q_{.05}(\alpha)$ values (i.e., the .05 quantile column) rather than the column of $q_{.95}(\alpha)$ values. For example, if $\hat{\alpha}_{LS3} = .80$, then $\hat{c}_U = 1.0$, since $\hat{\alpha}_{LS3} \geq q_{.05}(1)$.

## 5. ESTIMATION OF THE IMPULSE RESPONSE FUNCTION

The *impulse response function* of a time series $\{Y_t\colon t = 1, 2, \ldots\}$ measures the effect of a unit shock to $Y_t$ occurring at time $t$ (i.e., $U_t \to U_t + 1$ in Models 1–3) on the value of $Y_t$ at the future time periods $t + 1, t + 2, \ldots$. This function is of interest, because it quantifies the persistence of shocks to the time series and the latter is often of substantive importance. For example, much of the interest in the trend stationary/unit root debate for macroeconomic time series centers on the question of the degree of persistence of shocks to these series.

Given the definition of Models 1–3 in equations (2.1) and (2.2), it is clear that the effect of a unit change in $U_t$ on future $Y_t$ values is independent of $\mu$ and $\beta$. Hence, the impulse response functions of $\{Y_t\}$ in Models 1–3 are all equal and are equal to that of $\{Y_t^*\}$:

$$\text{(5.1)} \qquad \mathrm{IR}(h) = \alpha^h \quad \text{for} \quad h = 0, 1, 2, \ldots .$$

Two useful scalar measures of persistence that summarize the impulse response function are the *cumulative impulse response* (CIR) and the *half life of a unit shock* (HLS). The CIR gives the total cumulative effect of a unit shock on the entire future of the time series. For $\alpha \geq 0$, this HLS gives the length of time until the impulse response of a unit shock is half its original magnitude. For Models 1–3, we have

$$\text{(5.2)} \qquad \mathrm{CIR} = \sum_{h=0}^{\infty} \mathrm{IR}(h) = \frac{1}{1-\alpha} \quad \text{and} \quad \mathrm{HLS} = \left|\log(1/2)/\log(\alpha)\right|.$$

The impulse response function $\mathrm{IR}(h)$ for $h$ odd and the CIR for Models 1–3 are each strictly monotone functions of $\alpha$ for $\alpha \in (-1, 1]$. In consequence, median-unbiased estimators of each are obtained by plugging in the median-unbiased estimator $\hat{\alpha}_{Uj}$ in place of $\alpha$ in their definitions. Similarly, exact confidence intervals for $\mathrm{IR}(h)$ and CIR are given by $[(\hat{c}_L)^h, (\hat{c}_U)^h]$ for $h = 1, 3, 5, \ldots$ and $[1/(1 - \hat{c}_U),\ 1/(1 - \hat{c}_L)]$ respectively.

The function $\mathrm{IR}(h)$ for $h$ even is monotone in $\alpha$ for $\alpha$ in $[0, 1]$ or $(-1, 0]$. Hence, if the parameter space for $\alpha$ is given by $[0, 1]$ (or $[0, 1)$) or $(-1, 0]$, then median-unbiased estimators and exact confidence intervals for $\mathrm{IR}(h)$ for $h = 2, 4, 6, \ldots$ can be obtained as above but with $(\hat{\alpha}_{Uj}, \hat{c}_L, \hat{c}_U)$ replaced by $(\hat{\alpha}_{Uj} \vee 0,\ \hat{c}_L \vee 0,\ \hat{c}_U \vee 0)$ or $(\hat{\alpha}_{Uj} \wedge 0,\ \hat{c}_U \wedge 0)$, respectively, where $\vee$ and $\wedge$ denote the max and min operators respectively. The HLS is monotone in $\alpha$ for $\alpha$ in $[0, 1]$. Hence, if the parameter space for $\alpha$ is $[0, 1]$ or $[0, 1)$, then a median-unbiased estimator of HLS and an exact confidence interval for it are given by $\hat{HLS} =$

$|\log(1/2)/\log(\hat{\alpha}_{Uj} \vee 0)|$ and $[|\log(1/2)/\log(\hat{c}_L \vee 0)|, |\log(1/2)/\log(\hat{c}_U \vee 0)|]$ respectively.

## 6. MODEL SELECTION

The median-unbiased estimators introduced in Section 3 can be used to construct unbiased model selection procedures. By definition, a model selection procedure is *unbiased* if for any correct model the probability of selecting the correct model is at least as large as the probability of selecting each incorrect model. Unbiased selection procedures exhibit an intuitive impartiality property that may be useful if the selection of one model or another is a contentious issue. For example, in Model 2 or 3, one might want to select between the (trend) stationary model for which $\alpha \in (-1,1)$ and the unit root (with drift) model for which $\alpha = 1$. The concept of unbiased selection procedures is a special case of that of risk-unbiased decision rules; see Lehmann (1959, p. 12). If the loss function equals zero when the correct model is chosen and one otherwise, then a risk-unbiased decision rule for this problem is an unbiased selection procedure.

Given Model $j$ of Section 2 for $j = 1$, 2, or 3, suppose the problem is to select one of two models defined by $\alpha \in I_a$ and $\alpha \in I_b$, where $I_a$ and $I_b$ are intervals that partition the parameter space $(-1,1]$ for $\alpha$ ($(-1,1)$ in Model 1). For example, one might have $I_a = (-1,1)$ and $I_b = \{1\}$ or $I_a = (-1,.975)$ and $I_b = [.975,1]$. (The latter are considered in DeJong and Whiteman (1991) and Phillips (1991).)

The selection procedure we consider here is

(6.1) "*choose* $I_m$ *if* $\hat{\alpha}_{Uj} \in I_m$ *for* $m = a, b$,"

where $\hat{\alpha}_{Uj}$ is the median-unbiased estimator defined in Section 3. This procedure is unbiased. To see this, suppose $I_a$ lies to the left of $I_b$; then for all $\alpha \in I_a$, $P_\alpha(\hat{\alpha}_{Uj} \in I_b) \leq P_\alpha(\hat{\alpha}_{Uj} > \alpha) \leq 1/2 \leq P_\alpha(\hat{\alpha}_{Uj} \leq \alpha) \leq P_\alpha(\hat{\alpha}_{Uj} \in I_a)$, where the second and third inequalities use the median-unbiasedness of $\hat{\alpha}_{Uj}$. For $\alpha \in I_b$, the argument is analogous. We note that the selection procedure of (6.1) is also a valid level .5 (unbiased) test of $H_0$: $\alpha \in I_a$ versus $H_1$: $\alpha \in I_b$ and of $H_0$: $\alpha \in I_b$ versus $H_1$: $\alpha \in I_a$.

## 7. ROBUSTNESS TO NONNORMAL INNOVATIONS

In one respect, the methods introduced in Sections 2–6 do not rely on the assumption of normal innovations $\{U_t\colon t = 1, \ldots, T\}$. Any other scale family of distributions could be used in place of the normal family (provided new tables of quantiles of $\hat{\alpha}_{LSj}$ are constructed). On the other hand, the methods of Sections 2–6 do rely on the specification of *some* scale family of distributions of the innovations and it is of interest to know whether the methods are sensitive to the specification.

We consider the question: Suppose one assumes the innovations are normal but they actually have some other distribution, such as a $t$ distribution; then how close is the estimate $\hat{\alpha}_{Uj}$ to being median-unbiased and how close is the coverage probability of the confidence interval $[\hat{c}_L, \hat{c}_U]$ to the desired probability of $1-\alpha$? These questions are most easily answered by constructing analogues of Tables II and III that correspond to some nonnormal distributions of interest and then assessing the difference between Tables II and III and the new tables. This has been done (see Table IV) for the case of $t$ distributions with $d$ degrees of freedom for $d = 1, 2, 3$, and 10, shifted chi-square distributions with 4 and 8 degrees of freedom (i.e., chi-square distributions shifted left by 4 and 8 units, respectively, so that they have mean zero), and two mean zero discrete distributions with supports $\{-1, 1\}$ and $\{-3, 1\}$ and corresponding probabilities $(1/2, 1/2)$ and $(1/4, 3/4)$. Table IV also provides results for the case where the innovations are not iid, but rather, are ARCH random variables with parameter $\lambda = .3$ and $\lambda = .85$. That is, the distribution of $U_t$ conditional on $U_{t-1}, U_{t-2}, \ldots$ is $N(0, h_t)$, where $h_t = \xi + \lambda U_{t-1}^2$ and $\xi$ and $\lambda$ are positive constants. The distribution of $\hat{\alpha}_{LSj}$ is invariant with respect to $\xi$ and so $\xi$ is set equal to one without loss of generality. The ARCH innovations are uncorrelated, but conditionally heteroskedastic.

Table IV presents the .5, .05, and .95 quantiles of $\hat{\alpha}_{LS3}$ for the different innovation distributions considered. A corresponding table for $\hat{\alpha}_{LS2}$ is not reported here due to its similarity to Table IV (in terms of the magnitudes of the effects of nonnormality on the quantiles). Unlike Tables II and III, Table IV has been constructed using simulation. For each entry in the table, ten thousand repetitions were used. The standard errors on the simulated estimates of the .5, .05, and .95 quantiles are approximately .002, .001, and .001 respectively. For each repetition, an approximately stationary initial random variable $Y_0$ was obtained by setting $Y_{-201} = 0$ and generating $\{Y_t\colon t = -200, \ldots, 59\}$ according to Model 3 with $\mu = \beta = 0$ (without loss in generality) and with $\{U_t\colon t = -200, \ldots, 59\}$ being iid with the desired nonnormal distribution or ARCH with the parameter values specified above. The final sixty observations $\{Y_t\colon t = 0, \ldots, 59\}$ were used in simulating the quantiles of $\hat{\alpha}_{LS3}$.

The results for the various iid innovations can be summarized as follows. The median-unbiased estimates of $\alpha$ and the corresponding exact confidence intervals for $\alpha$ are quite robust against substantial skewness and kurtosis in the underlying innovation distribution. The estimates are somewhat more robust to nonnormality than are the confidence intervals. Only when the innovations have very thick tails (e.g., Cauchy distributions) are the estimates and confidence intervals in need of significant adjustment. The necessary adjustment in this case is not surprising, because the asymptotics that hold for normal innovations differ from those for Cauchy innovations. More surprising is the degree of robustness of the estimators and confidence intervals against finite variance symmetric and asymmetric innovation distributions.

Next, we discuss the results for the ARCH innovations. The estimator $\hat{\alpha}_{U3}$ (defined using $N(0, 1)$ innovations) is nearly median-unbiased when the innova-

TABLE IV

COMPARISON OF THE QUANTILES OF THE LEAST SQUARES ESTIMATOR OF $\alpha$ FOR DIFFERENT INNOVATION DISTRIBUTIONS FOR MODEL 3 AND SAMPLE SIZE 60

| $\alpha$/Distribution | $N(0,1)$ | $t_{10}$ | $t_3$ | $t_2$ | $t_1$ | $\chi_8^2-8$ | $\chi_4^2-4$ | $\{-1,1\}$ | $\{-3,1\}$ | ARCH(.3) | ARCH(.85) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | (a) .5 QUANTILE | | | | | | |
| .00 | −.034 | −.031 | −.033 | −.035 | −.032 | −.037 | −.038 | −.037 | −.038 | −.035 | −.035 |
| .20 | .155 | .159 | .158 | .156 | .161 | .153 | .152 | .151 | .150 | .153 | .149 |
| .40 | .344 | .348 | .348 | .345 | .353 | .343 | .340 | .340 | .340 | .340 | .334 |
| .60 | .532 | .535 | .535 | .534 | .544 | .530 | .530 | .529 | .528 | .527 | .518 |
| .80 | .715 | .717 | .719 | .718 | .730 | .714 | .712 | .713 | .713 | .710 | .702 |
| .90 | .799 | .802 | .803 | .803 | .819 | .799 | .797 | .798 | .797 | .795 | .786 |
| .95 | .834 | .837 | .838 | .839 | .858 | .833 | .834 | .833 | .832 | .831 | .820 |
| 1.00 | .853 | .855 | .856 | .859 | .873 | .851 | .852 | .852 | .852 | .852 | .842 |
| | | | | | (b) .05 QUANTILE | | | | | | |
| $\alpha$/Distribution | $N(0,1)$ | $t_{10}$ | $t_3$ | $t_2$ | $t_1$ | $\chi_8^2-8$ | $\chi_4^2-4$ | $\{-1,1\}$ | $\{-3,1\}$ | ARCH(.3) | ARCH(.85) |
| .00 | −.244 | −.244 | −.234 | −.226 | −.189 | −.242 | −.232 | −.249 | −.233 | −.286 | −.390 |
| .20 | −.062 | −.060 | −.054 | −.043 | −.001 | −.061 | −.048 | −.070 | −.051 | −.107 | −.221 |
| .40 | .126 | .127 | .136 | .145 | .190 | .128 | .140 | .118 | .136 | .081 | −.031 |
| .60 | .319 | .327 | .328 | .337 | .379 | .322 | .329 | .311 | .326 | .279 | .168 |
| .80 | .515 | .521 | .523 | .531 | .570 | .516 | .522 | .506 | .517 | .484 | .393 |
| .90 | .607 | .615 | .615 | .620 | .646 | .607 | .613 | .603 | .614 | .582 | .500 |
| .95 | .646 | .654 | .651 | .653 | .675 | .644 | .646 | .647 | .653 | .620 | .549 |
| 1.00 | .666 | .674 | .673 | .672 | .680 | .662 | .688 | .667 | .673 | .646 | .573 |
| | | | | | (c) .95 QUANTILE | | | | | | |
| $\alpha$/Distribution | $N(0,1)$ | $t_{10}$ | $t_3$ | $t_2$ | $t_1$ | $\chi_8^2-8$ | $\chi_4^2-4$ | $\{-1,1\}$ | $\{-3,1\}$ | ARCH(.3) | ARCH(.85) |
| .00 | .177 | .175 | .169 | .158 | .117 | .180 | .186 | .176 | .187 | .225 | .316 |
| .20 | .358 | .356 | .350 | .339 | .303 | .359 | .366 | .359 | .362 | .400 | .476 |
| .40 | .531 | .531 | .523 | .514 | .484 | .531 | .535 | .534 | .534 | .561 | .623 |
| .60 | .695 | .696 | .689 | .683 | .661 | .696 | .697 | .697 | .696 | .714 | .753 |
| .80 | .846 | .847 | .842 | .840 | .828 | .846 | .848 | .848 | .847 | .854 | .870 |
| .90 | .912 | .912 | .910 | .909 | .907 | .913 | .912 | .912 | .912 | .914 | .922 |
| .95 | .940 | .941 | .939 | .939 | .950 | .940 | .939 | .940 | .940 | .942 | .946 |
| 1.00 | .956 | .956 | .956 | .955 | .950 | .955 | .955 | .957 | .956 | .959 | .960 |

tions are ARCH(.3) and is slightly downward median-biased when the innovations are ARCH(.85). The dispersion of $\hat{\alpha}_{LS3}$ is greater when the innovations are ARCH(.3) or ARCH(.85) than when they are iid $N(0,1)$. In consequence, the 90% confidence intervals $[\hat{c}_L, \hat{c}_U]$ (defined using iid normal innovations) are shorter than they should be when the innovations are ARCH, especially ARCH(.85). In sum, ARCH innovations have a relatively small effect on the median-bias of the estimator $\hat{\alpha}_{U3}$, but they have a noticeable effect on the confidence intervals $[\hat{c}_L, \hat{c}_U]$, especially when the ARCH parameter $\lambda$ is large.

## 8. PROPERTIES OF THE LEAST SQUARES ESTIMATOR

In this section, we discuss various properties of the LS estimator of $\alpha$. See Hurwicz (1950), Orcutt and Winokur (1969), Fuller (1976), Phillips (1977), Dickey and Fuller (1981), Evans and Savin (1981, 1984), and Perron (1991) for further analysis of the finite sample properties of the LS estimator in Models 1 and 2.

We note first that conditional on the initial observation the LS estimator of $\alpha$ is both the maximum likelihood estimator and the Bayes estimator for an (improper) flat prior. Use of the latter estimator (in models with or without intercepts and time trends) has been advocated by Sims (1988, p. 467) and Sims and Uhlig (1991). Some of Sims' (1988) arguments for use of the flat-prior Bayes estimator, however, are based on the first-order AR model with no intercept or time trend and do not carry over to the more practically relevant models that contain an intercept or an intercept and time trend.

Tables I–III show the downward bias of the LS estimator of $\alpha$ in Models 1–3 when the sample size $T+1$ equals 60 and 100. The tables illustrate a striking difference between the bias for Model 1 with no intercept and the biases for Models 2 and 3. For Model 1 the bias is very small, being roughly $-.01$ or less for most values of $\alpha$. Only if $\alpha$ is very close to one is a bias of this magnitude of import. Thus, for Model 1, it often does not matter appreciably if one uses the LS estimator or the median-unbiased estimator. On the other hand, the downward biases for Models 2 and 3 are quite large, especially for $\alpha$ large and for Model 3. In these Models, there *is* an appreciable difference between the LS estimator and the median-unbiased estimator.

Tables I–III also show the variability of the LS estimator of $\alpha$ as measured by its 90% range for Models 1–3. For Model 1, the variability of the LS estimator is a decreasing function of $\alpha$ and it decreases dramatically as $\alpha$ approaches one. Sims (1988, p. 469) has pointed out this property of the LS estimator and argues that it counter-balances the downward bias of the LS estimator with the net result being that one is as likely to get a spuriously high LS estimate of $\alpha$ when $\alpha$ is less than 1 as one is to get a spuriously low LS estimate of $\alpha$ when $\alpha$ is close to one. Indeed, for Model 1 with $T+1=60$ we find that when $\alpha=.999$ the .05 quantile of the LS estimator is .94 and when $\alpha=.94$ the .95 quantile of the LS estimator is .98, so there is a near balance between spuriously high and spuriously low estimates.

For Models 2 and 3, the variability of the LS estimator is also a decreasing function of $\alpha$. It is not, however, nearly as sensitive to the value of $\alpha$ as in Model 1. For example, the 90% range for $\alpha = .8$ is only 1.14 times as large as for $\alpha = .999$ in Model 3 and only 1.36 times as large in Model 2, whereas it is 3.9 times as large in Model 1. This relative insensitivity to $\alpha$ of the variability of the LS estimator in Models 2 and 3 coupled with its large downward biases in these models, causes the LS estimator to yield spuriously low estimates of $\alpha$ when $\alpha$ is near 1 with much greater frequency than spuriously high estimates of $\alpha$ when $\alpha$ is less than 1. For example, when $\alpha = 1$, the .05 quantile of $\hat{\alpha}_{LS3}$ is .67 for $T + 1 = 60$, whereas when $\alpha = .67$, the .95 quantile of $\hat{\alpha}_{LS3}$ is only .75 (rather than 1).

Next, we discuss an argument given by Sims (1988, p. 469) regarding the unbiasedness of Bayes estimators. The argument given is that the mean of a Bayes estimator equals the mean of the random parameter being estimated. This might be referred to as *unconditional* (on the parameter) *unbiasedness* of the Bayes estimator. The above argument is subject to a number of caveats, however, and at best illustrates only a very weak sort of unbiasedness that does not reflect impartiality of the estimator in our opinion.

First, the argument does not apply to Bayes estimators based on improper priors.[6] In particular, it does not apply to the LS estimator. The LS estimator is downward mean-biased conditionally on $\alpha$ for each $\alpha$ in $(0, 1]$ in Models 1–3 and, hence, is downward biased unconditionally for any distribution of $\alpha$ with support in $(0, 1]$.

Second, the argument relies on the prior that is used to form the Bayes estimator being equal to the marginal distribution of the parameter. If they differ, unconditional unbiasedness does not hold.

Third, the criterion of unconditional unbiasedness is very weak. To see why, suppose the parameter $\alpha$ has distribution $P(\alpha)$ on $(-1, 1]$ with mean equal to .8. Then, the Bayes estimator with prior $P(\alpha)$ has unconditional mean .8 and is unconditionally unbiased. But, the constant estimator .8 also has mean .8 and is unconditionally unbiased. This constant estimator does not exhibit any sort of impartiality—it is completely partial. Thus, the property of unconditional unbiasedness does not reflect impartiality of an estimator.

## 9. EXAMPLES

In this section, we apply the methods introduced above to a number of data series that have been analyzed in the literature. First, we consider eight real exchange rate series that have been investigated recently using Bayesian meth-

[6] The argument for the unconditional unbiasedness of a Bayes estimator is as follows: Let $\hat{\alpha}_B(Y)$ denote the Bayes estimator of $\alpha$ as a function of the data $Y$. Then, $E_Y \hat{\alpha}_B(Y) = E_Y E(\alpha | Y) = E_{Y,\alpha}\alpha = E_\alpha \alpha$, where $E_{Y,\alpha}$, $E_Y$, and $E_\alpha$ denote expectations with respect to $(Y, \alpha)$, $Y$, and $\alpha$ respectively. This argument relies on $E_\alpha|\alpha|$ being finite, which does not hold if the prior on $\alpha$ is improper. In fact, the posterior mean of $\alpha$ does not necessarily exist if the prior on $\alpha$ is improper, e.g., it does not exist for a Jeffrey's prior; see Phillips (1991).

ods by Schotman and van Dijk (1991a). Model 2 is used for these series (as in Schotman and Dijk (1991a)).[7] Second, we consider two of the series considered by Nelson and Plosser (1982), viz., the velocity of money and industrial production in the United States. Model 3 is used for these series. (Nelson and Plosser (1982) use Model 3 for the velocity series, but allow for an AR(6) model with intercept and time trend for the industrial production series. We argue that the much more parsimonious Model 3 may provide an adequate representation for industrial production.[8])

The first six real exchange rate series are of the U.S. dollar against the currencies of France (FR), West Germany (WG), Japan (JP), Canada (CA), United Kingdom (UK), and the Netherlands (NL). The last two series are of the German Dmark against the currencies of France and the Netherlands, both of which are fellow members of the European Monetary System. The real exchange rate series is defined to be $Y_t = e_t - P_t + P_t^*$, where $e_t$ is the logarithm of the end of month nominal exchange rate expressed as the domestic price of one unit of foreign currency and $P_t$ and $P_t^*$ are the logarithms of the monthly consumer price indices of the domestic and foreign countries respectively. The nominal exchange rate and consumer price indices are from the IFS databank—series ae and 64—for the sample period 73:01 to 88:07, which yields a sample size $T+1$ of 187, except for Canada and the UK where the terminal date is 88:06 and $T+1$ is 186.

The velocity and industrial production series are annual series covering 1869–1970 and 1860–1970 respectively. Hence, the sample sizes $T+1$ of these series $\{Y_t: t = 0, \ldots, T\}$ are 102 and 111. Both series are from data files supplied by Nelson and Plosser. The velocity series is originally from Friedman and Schwartz (1963) with revisions provided by Schwartz. The industrial production series is from *Long Term Economic Growth* (1973).

Table V provides the results for the real exchange rate series. For each of the U.S. dollar exchange rates except UK/US, the median-unbiased estimate of $\alpha$ is 1.00. For UK/US it is .995. For the FR/WG and NL/WG exchange rates, the median-unbiased estimates of $\alpha$ are .968 and .965 respectively. In each case, the median-unbiased estimates of $\alpha$ exceed those of the LS estimates by between .017 and .022. While the magnitudes of these differences are small in absolute terms, they are large in terms of their implications for the persistence of the time series. This is illustrated by the corresponding median-unbiased and LS estimates of the impulse response function, CIR, and HLS given in Table V.

[7] Schotman and van Dijk (1991a) use Model 2 but condition on $Y_0$. We view it to be an attribute of our procedure that we do not need to condition on $Y_0$, since $Y_0$ contains information about $\alpha$ when $\alpha \in (-1, 1)$.

[8] The LS estimates of the five regressors $\Delta Y_{t-1}, \ldots, \Delta Y_{t-5}$ that appear in Nelson and Plosser's (1982) model for industrial production but do not appear in Model 3, equation (2.3), are .10, −.04, .04, −.05, −.22, respectively, each with a standard error estimate of .10. This pattern of coefficients seems somewhat implausible. In fact, the $F$ statistic for testing jointly whether the coefficients on all five $\Delta Y_{t-j}$ regressors are zero is .2, which is very small (standard asymptotics are applicable for this test statistic whether or not $\alpha = 1$). This and other evidence suggest that Model 3 with no $\Delta Y_{t-j}$ regressors may be appropriate for modelling the industrial production series.

TABLE V

REAL EXCHANGE RATE SERIES[a]

| Data Series | Estimator | $\alpha$ | IR(3) | IR(12) | IR(36) | CIR | HLS |
|---|---|---|---|---|---|---|---|
| FR/US | LS | .982 | .947 | .804 | .520 | 55.6 | 38.2 |
| | Median- | 1.00 | 1.00 | 1.00 | 1.00 | $\infty$ | $\infty$ |
| | Unbiased | [.97, 1.0] | [.91, 1.0] | [.69, 1.0] | [.33, 1.0] | [33.3, $\infty$] | [22.8, $\infty$] |
| WG/US | LS | .982 | .947 | .804 | .520 | 55.6 | 38.2 |
| | Median- | 1.00 | 1.00 | 1.00 | 1.00 | $\infty$ | $\infty$ |
| | Unbiased | [.97, 1.0] | [.91, 1.0] | [.69, 1.0] | [.33, 1.0] | [33.3, $\infty$] | [22.8, $\infty$] |
| JP/US | LS | .983 | .950 | .814 | .539 | 58.8 | 40.4 |
| | Median- | 1.00 | 1.00 | 1.00 | 1.00 | $\infty$ | $\infty$ |
| | Unbiased | [.97, 1.0] | [.91, 1.0] | [.69, 1.0] | [.33, 1.0] | [33.3, $\infty$] | [22.8, $\infty$] |
| CA/US | LS | .983 | .950 | .814 | .539 | 58.8 | 40.4 |
| | Median- | 1.00 | 1.00 | 1.00 | 1.00 | $\infty$ | $\infty$ |
| | Unbiased | [.97, 1.0] | [.91, 1.0] | [.69, 1.0] | [.33, 1.0] | [33.3, $\infty$] | [22.8, $\infty$] |
| UK/US | LS | .973 | .921 | .720 | .373 | 37.0 | 25.3 |
| | Median- | .995 | .985 | .942 | .835 | 200 | 138 |
| | Unbiased | [.96, 1.0] | [.88, 1.0] | [.61, 1.0] | [.23, 1.0] | [25, $\infty$] | [22.8, $\infty$] |
| NL/US | LS | .981 | .944 | .794 | .501 | 52.6 | 36.1 |
| | Median- | 1.00 | 1.00 | 1.00 | 1.00 | $\infty$ | $\infty$ |
| | Unbiased | [.97, 1.0] | [.91, 1.0] | [.69, 1.0] | [.33, 1.0] | [33.3, $\infty$] | [22.8, $\infty$] |
| FR/WG | LS | .950 | .857 | .540 | .158 | 20.0 | 13.5 |
| | Median- | .969 | .910 | .685 | .322 | 32.3 | 22.0 |
| | Unbiased | [.93, 1.0] | [.80, 1.0] | [.42, 1.0] | [.07, 1.0] | [14.3, $\infty$] | [9.6, $\infty$] |
| NL/WG | LS | .947 | .849 | .520 | .141 | 18.9 | 12.7 |
| | Median- | .966 | .901 | .660 | .288 | 29.4 | 20.0 |
| | Unbiased | [.92, 1.0] | [.78, 1.0] | [.37, 1.0] | [.05, 1.0] | [12.5, $\infty$] | [8.3, $\infty$] |

[a] The results of this table are based on Model 2 of Section 2. IR($h$) denotes the impulse response function at time horizon $h$. CIR denotes the cumulative impulse response. The entries in the rows labeled LS are the estimates of $\alpha$, IR($h$), and CIR obtained using the least squares estimates of $\alpha$ from the regression equation (2.3) Model 2. The entries in the rows labeled median-unbiased are the median-unbiased estimates $\hat{\alpha}_{Uj}$, $\hat{\text{IR}}(h)$, and $\hat{\text{CIR}}$ of $\alpha$, IR($h$), and CIR, respectively, defined in Sections 3 and 5. The intervals in square brackets below the median-unbiased estimates are the 90% central (i.e., $p_1 = p_2 = .05$) confidence intervals introduced in Sections 4 and 5.

The median-unbiased estimates and exact confidence intervals given in this table were determined using quantiles generated by simulation (with 10,000 repetitions) for the exact sample size of each series. Interpolation of the quantiles given in Table II for sample sizes 150 and 200 produces nearly the same results.

The residuals $\{Y_t - \hat{\alpha}_{U2}Y_{t-1}\colon\ t = 1, \ldots, T\}$ for the eight real exchange rate series reveal evidence of nonnormality for some of the series. The robustness of the results in Table V to nonnormality has been checked by calculating the median-unbiased estimates of $\alpha$ and the corresponding exact 90% confidence intervals for $\alpha$ for innovations given by a number of $t$ and shifted $\chi^2$ distributions whose degrees of freedom yield coefficients of skewness and kurtosis that cover the range found in the residual series. The following distributions were considered: $t_3, t_5, t_{11}, t_{14}, \chi^2_4 - 4$, and $\chi^2_8 - 8$. In short, the results reported in Table V were hardly changed at all when any of the above distributions was used. The median-unbiased estimates of $\alpha$ were unchanged except for the FR/WG and NL/WG series for which the estimates differed by at most .001 and .002. The confidence intervals were unchanged.

The robustness of the results of Table V to first order ARCH innovations also has been checked. For each data series, Model 2 was estimated with ARCH innovations by maximum likelihood, the estimated ARCH parameter was taken

TABLE VI

VELOCITY AND INDUSTRIAL PRODUCTION SERIES[a]

| Data Series | Estimator | $\alpha$ | IR(2) | IR(4) | IR(16) | CIR | HLS |
|---|---|---|---|---|---|---|---|
| Velocity | LS | .94 | .88 | .78 | .37 | 16.7 | 11.2 |
| | Median-Unbiased | 1.0 [.93, 1.0] | 1.0 [.86, 1.0] | 1.0 [.75, 1.0] | 1.0 [.31, 1.0] | $\infty$ [14.3, $\infty$] | $\infty$ [9.6, $\infty$] |
| Industrial Production | LS | .84 | .71 | .50 | .06 | 6.3 | 4.0 |
| | Median-Unbiased | .89 [.79, 1.0] | .79 [.62, 1.0] | .63 [.39, 1.0] | .15 [.02, 1.0] | 9.1 [4.8, $\infty$] | 5.9 [2.9, $\infty$] |

| | | $\tilde{\mu}$ | $\tilde{\beta}$ | $\sigma^2$ | Durbin-Watson | Skewness | Excess Kurtosis |
|---|---|---|---|---|---|---|---|
| Velocity | LS | .052 (.035) | −.00032 (.00049) | .0045 | 1.78 | −.49 | .35 |
| | Median-Unbiased | −.033 (.014) | .00040 (.00023) | .0046 | 1.82 | −.35 | .16 |
| Industrial Production | LS | .043 (.019) | .0066 (.0022) | .0096 | 1.79 | −1.01 | 1.62 |
| | Median-Unbiased | .044 (.019) | .0045 (.00029) | .0096 | 1.86 | −.96 | 1.33 |

[a] The top half of Table VI is analogous to Table V except that Model 3 equation (2.3) is estimated rather than Model 2 equation (2.3). The bottom half of Table VI provides additional information. The entries in the rows labeled LS give the LS estimates of the parameters $\tilde{\mu}$, $\tilde{\beta}$, and $\sigma^2$ (with standard error estimates in parentheses) as well as the Durbin-Watson statistic and the sample coefficients of skewness and excess kurtosis from the regression of $Y_t$ on $(1, t, Y_{t-1})$. The rows labeled median-unbiased report the corresponding statistics from the regression with $\alpha$ restricted to equal its median-unbiased estimate, i.e., the regression of $Y_t - \hat{\alpha}_{U3} Y_{t-1}$ on $(1, t)$.

to be the truth, and median-unbiased estimates of $\alpha$ and corresponding exact 90% confidence intervals for $\alpha$ based on ARCH innovations were calculated with the estimated ARCH parameter. The ML estimates of the ARCH parameters for the eight series, as ordered in Tables V, are .08, .10, .16, .09, .05, .12, .23, and .11. The ARCH-based results are the same as those of Table V except for the FR/WG series for which the estimate and 90% CI for $\alpha$ are .970 and [.92, 1.00] rather than .969 and [.93, 1.00]. Thus, the results of Table V seem to be quite robust against empirically relevant amounts of conditional heteroskedasticity and nonnormality of the innovations.

Schotman and van Dijk's (1991a) Bayesian estimates, based on a data-dependent uniform prior for $\alpha$ on $[a^*, 1)$ and no prior mass on $\alpha = 1$, lie between the LS and median-unbiased estimates. For the series ordered as above and as in Table V, their estimates are .982, .982, .990, .987, .977, .983, .954, and .965. If prior probability mass is placed on $\alpha = 1$, then their estimates are a corresponding weighted average of the estimates listed above and one. Depending on how much prior mass is placed on $\alpha = 1$, Schotman and van Dijk's estimates may be close to or somewhat distant from the median-unbiased estimates. An advantage of the median-unbiased estimates is that they do not rely on the contentious choice of how much prior mass is placed on $\alpha = 1$.

Table VI provides the results for the velocity and industrial production series. For velocity, the median-unbiased estimate of $\alpha$ is 1.0 compared to the LS estimate of .94. In fact, the LS estimate is barely above the bottom edge of the

90% central confidence interval of [.93, 1.0]. In consequence, this series exhibits the most dramatic differences of all the series between the LS and median-unbiased estimated persistence of shocks to the series, as measured by IR($h$), CIR, and HLS. The differences are quite substantial. The unbiased model selection procedure of Section 6 chooses the random walk with drift model over the trend stationary model.

For industrial production, the median-unbiased estimate of $\alpha$ is .89 as compared to the LS estimate of .84. This yields a 50% larger estimate of CIR and HLS using the unbiased estimate rather than the LS estimate. This is a fairly large difference, but it is not as dramatic as for the velocity series. The length of the 90% central confidence interval [.79, 1.0] for $\alpha$ indicates that $\alpha$ cannot be estimated very precisely in this example. Although the random walk with drift model is not rejected at level .05 (since the upper bound one-sided confidence interval for $\alpha$ contains 1), the unbiased model selection procedure of Section 6 chooses the trend stationary model.

The difference between the median-unbiased estimates, confidence intervals, and unbiased model selection results for the velocity series and those for the industrial production series indicates the value of having such statistics available. The outcome of a level .05 test of a unit root is the same for both series—it does not reject. But the evidence for a unit root is much stronger for the velocity series than for the industrial production series.

*Cowles Foundation for Research in Economics, Yale University, Box 2125, Yale Station, New Haven, CT 06520, U.S.A.*

*Manuscript received October, 1991; final revision received July, 1992.*

## APPENDIX A

We now prove the invariance property of $\hat{\alpha}_{LSj}$ that is stated in Section 2. First, consider Model 3. To obtain $\hat{\alpha}_{LS3}$, first regress $Y_t$ on $(1, t)$ and $Y_{t-1}$ on $(1, t)$ for $t = 1, \ldots, T$. Then, $\hat{\alpha}_{LS3}$ equals the LS estimate from the regression of the residuals from the former regression on the residuals from the latter regression. Since $Y_t = \mu + \beta t + Y_t^*$ by (2.2), both sets of residuals are invariant with respect to the values of $\mu$ and $\beta$. In consequence, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to $(\mu, \beta)$.

Given this invariance, we can suppose $\mu = \beta = 0$ when analyzing the possible dependence of $\hat{\alpha}_{LS3}$ on $\sigma^2$ and of $\hat{\alpha}_{LS3}$ on $Y_0$ when $\alpha = 1$. When $\mu = \beta = 0$, $Y_t = Y_t^*$ for $t = 0, \ldots, T$. Multiplying $\sigma^2$ by a positive constant $c$ in (2.1) causes $Y_t^*$ and $Y_t$ to be multiplied by the same constant $c$ for $t = 0, \ldots, T$ when $\alpha \in (-1, 1)$. In consequence, the residuals from the regression of $Y_t$ on $(1, t)$ and of $Y_{t-1}$ on $(1, t)$ are multiplied by the same constant. This constant cancels out when the former residuals are regressed on the latter to obtain $\hat{\alpha}_{LS3}$. Thus, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to $\sigma^2$ when $\alpha \in (-1, 1)$.

Next, suppose $\mu = \beta = 0$ and $\alpha = 1$. In this case, $Y_t = Y_t^* = Y_0 + \Sigma_{s=1}^t U_s$ for $t = 0, \ldots, T$ and the residuals from the regression of $Y_t$ on $(1, t)$ and of $Y_{t-1}$ on $(1, t)$ are invariant with respect to the value of $Y_0$. In consequence, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to the value of $Y_0$ when $\alpha = 1$. Given this invariance, suppose $Y_0 = 0$. Then, the multiplication of $\sigma^2$ by a constant $c$ causes $Y_t^*$ and $Y_t$ to be scaled up by the same constant. As above, this leaves $\hat{\alpha}_{LS3}$ unchanged. Hence, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to $\sigma^2$ when $\alpha = 1$. This completes the proof for Model 3.

The proofs for Models 1 and 2 are analogous, except that for Model 1 the argument for the invariance of the LS estimator with respect to $Y_0$ when $\mu = \beta = 0$ and $\alpha = 1$ does not hold (because

$Y_t$ and $Y_{t-1}$ are regressed on 0 rather than on $(1, t)$ in this case). This does not affect the result stated in (2.4) because $\alpha = 1$ is excluded in Model. 1.

## APPENDIX B

Here we describe the numerical procedure that was used to compute the .05, .5, and .95 quantiles of the LS estimators $\hat{\alpha}_{LSj}$, $j = 1,2,3$ which are reported in Tables I–III. The basic idea is to write $P_\alpha(\hat{\alpha}_{LSj} \le c)$ as the probability that a quadratic form in standard normal variates is less than or equal to zero, to apply the Imhof (1961) algorithm to compute such probabilities, and to compute the desired quantiles using an iterative procedure that involves computing such probabilities for different values of $c$ until the value of $c$ is found that yields the desired probability .05, .5, or .95.

First, we define the LS estimator $\hat{\alpha}_{LSj}$ as

$$\text{(B.1)} \qquad \alpha_{LSj} = \underset{\sim}{Y}' D_0' (I - P_j) D_T \underset{\sim}{Y} / \underset{\sim}{Y}' D_T' (I - P_j) D_T \underset{\sim}{Y} \quad \text{for} \quad j = 1,2,3,$$

where $\underset{\sim}{Y} = (Y_0, \ldots, Y_T)'$, $D_0 = [\underset{\sim}{0} : I] \in R^{T \times (T+1)}$, $D_T = [I : \underset{\sim}{0}] \in R^{T \times (T+1)}$, $I$ equals the $T$ dimensional identity matrix, $P_1 = \underset{\sim}{0} \in R^{T \times T}$, $P_j = X_j (X_j' X_j)^{-1} X_j'$ for $j = 2,3$, $X_1 = (1, \ldots, 1)' \in R^T$, $X_2 = (X_1 : x_2) \in R^{T \times 2}$, and $x_2 = (1, 2, \ldots, T)'$.

Next, we write $\underset{\sim}{Y}$ in terms of the underlying errors $\{U_t\}$. By the invariance properties of $\hat{\alpha}_{LSj}$ described in Section 2, it suffices to consider the case where $\mu = \beta = 0$ and $\sigma^2 = 1$ in Models 1–3 and $Y_0 = 0$ when $\alpha = 1$ in Models 2 and 3. In this case, we have

$$\underset{\sim}{Y} = R_\alpha \underset{\sim}{U}, \quad \text{where} \quad \underset{\sim}{U} = (U_0, \ldots, U_T)', \quad \underset{\sim}{U} \sim N(\underset{\sim}{0}, I_{T+1}),$$

$$\text{(B.2)} \qquad R_\alpha = \begin{bmatrix} b & 0 & 0 & \cdots & 0 & 0 \\ b\alpha & 1 & 0 & \cdots & 0 & 0 \\ b\alpha^2 & \alpha & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b\alpha^T & \alpha^{T-1} & \alpha^{T-2} & \cdots & \alpha & 1 \end{bmatrix},$$

$$b = 1/(1 - \alpha^2)^{1/2} \quad \text{when} \quad \alpha \in (-1,1), \quad \text{and} \quad b = 0 \quad \text{otherwise.}$$

We now can write

$$\text{(B.3)} \qquad P_\alpha(\hat{\alpha}_{LSj} \le c) = P(\underset{\sim}{U}' W_{\alpha jc} \underset{\sim}{U} \le 0), \quad \text{where}$$

$$W_{\alpha jc} = R_\alpha' \left[ D_0' (I - P_j) D_T / 2 + D_T' (I - P_j) D_0 / 2 - c D_T' (I - P_j) D_T \right] R_\alpha.$$

(Note that, as defined, the weight matrix $W_{\alpha jc}$ is symmetric.) Since $\underset{\sim}{U}' W_{\alpha jc} \underset{\sim}{U}$ is a quadratic form in standard normal variates, $P_\alpha(\hat{\alpha}_{LSj} \le c)$ can be computed using Imhof's (1961) algorithm. This was done by employing the FORTRAN subroutines given by Koerts and Abrahamse (1971, Ch. 9). In their subroutine FQUAD, the tolerance parameters EPS1 and EPS2 were set equal to $10^{-6}$ and $10^{-36}$ respectively. EPS1 is an upper bound on the truncation error in computing the integral. EPS2 is an upper bound on the effect of halving the grid size used to calculate the truncated integral via Simpson's rule.

To obtain the $p$th quantile of $\hat{\alpha}_{LSj}$, we need to find the value $c_p$ such that $P_\alpha(\hat{\alpha}_{LSj} \le c_p) = p$. Since we can compute $P_\alpha(\hat{\alpha}_{LSj} \le c)$ for arbitrary $c$ and $P_\alpha(\hat{\alpha}_{LSj} \le c)$ is increasing in $c$, a simple iterative procedure was used to determine $c_p$. The results given in Tables I–III were obtained in this manner.

We note that two alternative methods of calculating the quantiles of $\hat{\alpha}_{LSj}$ are numerical complex integration and simulation. These methods are particularly useful when the sample size is large—say, greater than one hundred, and especially when it is greater than two hundred.

## REFERENCES

ANDREWS, D. W. K. (1989): "Power in Econometric Applications," *Econometrica*, 57, 1059–1090.

ANDREWS, D. W. K., AND H. Y. CHEN (1992): "Approximately Median-unbiased Estimation of Autoregressive Models with Application to U.S. Macroeconomic and Financial Time Series," Cowles Foundation Discussion Paper No. 1026, Yale University.

ANDREWS, D. W. K., AND P. C. B. PHILLIPS (1987): "Best Median-Unbiased Estimation in Linear Regression with Bounded Asymmetric Loss Functions," *Journal of the American Statistical Association*, 82, 886–893.

BHARGAVA, A. (1986): "On the Theory of Testing for Unit Roots in Observed Time Series," *Review of Economic Studies*, 53, 369–384.

DEJONG, D. N., J. C. NANKERVIS, N. E. SAVIN, AND C. H. WHITEMAN (1992a): "Integration versus Trend-stationarity in Time Series," *Econometrica*, 60, 423–433.

——— (1992b): "The Power Problems of Unit Root Tests in Time Series with Autoregressive Errors," *Journal of Econometrics*, 53, 323–343.

DEJONG, D., AND C. H. WHITEMAN (1991): "Reconsidering Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 28, 221–254.

DICKEY, D. A., AND W. A. FULLER (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

——— (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

DUFOUR, J.-M. (1990): "Exact Tests and Confidence Sets in Linear Regressions with Autocorrelated Errors," *Econometrica*, 58, 475–494.

EVANS, G. B. A., AND N. E. SAVIN (1981): "Testing for Unit Roots: 1," *Econometrica*, 49, 753–779.

——— (1984): "Testing for Unit Roots: 2," *Econometrica*, 52, 1241–1269.

FRIEDMAN, M., AND A. J. SCHWARTZ (1963): *A Monetary History of the United States, 1867–1960*. Princeton, NJ: Princeton University Press.

FULLER, W. A. (1976): *Introduction to Statistical Time Series*. New York: John Wiley and Sons.

HURWICZ, L. (1950): "Least-squares Bias in Time Series," in *Statistical Inference in Dynamic Economic Models*, ed. by T. C. Koopmans. Wiley: New York.

IMHOF, J. P. (1961): "Computing the Distribution of Quadratic Forms in Normal Variates," *Biometrika*, 48, 419–426.

KENDALL, M. G. (1954): "Note on Bias in the Estimation of Autocorrelation," *Biometrika*, 41, 403–404.

KIVIET, J. F., AND G. D. A. PHILLIPS (1990): "Exact Similar Tests for the Root of a First-Order Autoregressive Regression Model," unpublished manuscript, Department of Actuarial Science and Econometrics, University of Amsterdam.

KOERTS, J., AND A. P. J. ABRAHAMSE (1971): *On the Theory and Application of the General Linear Model*. Rotterdam: Rotterdam University Press.

LE BRETON, A., AND D. T. PHAM (1989): "On the Bias of the Least Squares Estimator for the First Order Autoregressive Process," *Annals of the Institute of Statistical Mathematics*, 41, 555–563.

LEHMANN, E. L. (1959): *Testing Statistical Hypotheses*. New York: John Wiley and Sons.

*Long Term Economic Growth* (1973): Washington, D.C.: U.S. Department of Commerce.

MARRIOTT, F. H. C., AND J. A. POPE (1954): "Bias in the Estimation of Autocorrelations," *Biometrika*, 41, 390–402.

NELSON, C. R., AND C. I. PLOSSER (1982): "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 10, 139–162.

ORCUTT, G. H., AND H. S. WINOKUR, JR. (1969): "First Order Autoregression: Inference, Estimation, and Prediction," *Econometrica*, 37, 1–14.

PERRON, P. (1991): "A Continuous Time Approximation to the Unstable First-Order Autoregressive Process: The Case without an Intercept," *Econometrica*, 59, 211–239.

PHILLIPS, P. C. B. (1977): "Approximations to Some Finite Sample Distributions Associated with a First-Order Stochastic Difference Equation," *Econometrica*, 45, 463–485.

——— (1991): "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends," *Journal of Applied Econometrics*, 6, 333–364.

QUENOUILLE, M. H. (1949): "Approximate Tests of Correlation in Time-Series," *Journal of the Royal Statistical Society Series B*, 11, 68–84.

——— (1956): "Notes on Bias in Estimation," *Biometrika*, 43, 353–360.

RUDEBUSCH, G. D. (1992): "Trends and Random Walks in Macroeconomic Time Series: A Re-examination," *International Economic Review*, 33, forthcoming.

SCHOTMAN, P., AND H. K. VAN DIJK (1991a): "A Bayesian Analysis of the Unit Root in Real Exchange Rates," *Journal of Econometrics*, 49, 195–238.

——— (1991b): "Posterior Analysis of Possibly Integrated Time Series with an Application to Real GNP," in *New Directions in Time Series Analysis, IMA Volume in Mathematics and its Applications*, ed. by E. Parzen et al. New York: Springer Verlag, forthcoming.

SHAMAN, P., AND R. A. STINE (1988): "The Bias of Autoregressive Coefficient Estimators," *Journal of the American Statistical Association*, 83, 842–848.

SIMS, C. A. (1988): "Bayesian Skepticism on Unit Root Econometrics," *Journal of Economic Dynamics and Control*, 12, 463–474.

SIMS, C. A., AND H. UHLIG (1991): "Understanding Unit Rooters: A Helicopter Tour," *Econometrica*, 59, 1591–1599.

STINE, R. A., AND P. SHAMAN (1989): "A Fixed Point Characterization for Bias of Autoregressive Estimators," *Annals of Statistics*, 17, 1275–1284.

STOCK, J. H. (1991): "Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series," *Journal of Monetary Economics*, 28, 435–459.